

Category theoretic properties of the A. Rényi and C. Tsallis entropies.

György Steinbrecher

Physics Department, University of Craiova, A. I. Cuza 13, 200585
Craiova, Romania. Email: gyorgy.steinbrecher@gmail.com

Alberto Sonnino

Karlsruhe Institute of Technologies (KIT)
Department of Electrical Engineering and Information Technologies
D-76131 Karlsruhe, Germany &
Ecole Polytechnique de Louvain (EPL)
Université Catholique de Louvain (UCL)
Rue Archimede 1 bte L6.11.01, 1348 Louvain-la-Neuve - Belgium.
Email: alberto.sonnino@gmail.com

Giorgio Sonnino

Department of Theoretical Physics and Mathematics
Université Libre de Bruxelles (ULB)
Campus Plaine CP 231, Bvd de Triomphe, 1050 Brussels, Belgium.
& Royal Military School (RMS)
Av. de la Renaissance 30 1000 Brussels - Belgium
Email: gsonnino@ulb.ac.be

December 21, 2015

Abstract

The problem of embedding the Tsallis and Rényi entropies in the framework of category theory and their axiomatic foundation is studied. To this end, we construct a special category MES related to measured spaces. We prove that both of the Rényi and Tsallis entropies can be imbedded in the formalism of category theory by proving that the same basic functional that appears in their definitions, as well as in the associated Lebesgue space norms, has good algebraic compatibility properties. We prove that this functional is both additive and multiplicative with respect to the direct product and the disjoint sum (the coproduct) in the category MES, so it is a natural candidate for the measure of information or uncertainty. We prove that the category MES can be extended to monoidal category, both with respect to the direct product as well as

to the coproduct. The basic axioms of the original Rényi entropy theory are generalized and reformulated in the framework of category MES and we prove that these axioms foresee the existence of an universal exponent having the same values for all the objects of the category MES. In addition, this universal exponent is the parameter, which appears in the definition of the Tsallis and Rényi entropies.

1 Introduction

The discovery of two related generalizations of the classical Shannon entropy [1] is a remarkable coincidence in the history of abstract probability theory and statistical physics. A. Rényi introduced a possible generalization [2] of the classical Shannon entropy by pure axiomatic extension of the Fadeev axioms [3], [4] that define uniquely the Shannon entropy. On the other hand, the generalized entropy [5], [6] introduced by C. Tsallis was useful to extend the classical maximum entropy principle such that the heavy tailed distributions observed in a large scale of physical processes [7], [8], [9], [10], could be derived from (generalized) maximum entropy principles. The interest in the study of the generalizations of the Shannon entropy in the recent years is due to the multiple applications of the Tsallis and Rényi entropy or the associated Rényi divergence [7], [8], [11], [12]. We mention also that similar to the classical *H theorem of L. Boltzmann*, the generalizations of the Rényi entropy, as well as the original Rényi entropy, is a Liapunov functional for a large class of stochastic processes described by generalized Fokker-Planck equations, more exactly by Fokker-Planck equation where the drift term and the diffusion tensor are itself dependent on some external random variable [13]. We mention that in the case of suitable singular limiting procedure, both the Tsallis and Rényi entropies give the same limit: the Shannon entropy. The Rényi entropy is additive while the Tsallis entropy is not. Despite the Rényi and Tsallis entropies give the same results in the case of problems associated to the determination of the probability density function from the Maximum Entropy principles, because they are algebraically related by simple formulae, the non-additivity of the Tsallis entropy generated many discussions in the physical literature. On the other hand, by formulating the basic axioms [2], A. Rényi introduced new concepts (*incomplete random variables and incomplete distributions*) that are not included in the standard terminology of the probability theory. Also the formulation of the *Postulate 5'* [2], is not the simplest, mathematically natural.

In this work we develop a formalism in the framework of the category theory [14] and [15] for the study of generalized entropies. The category theory is the branch of mathematics that plays a central role in the logical foundation and synthesis of the whole contemporary mathematics. In particular, the category theory allows avoiding the paradoxes of the classical set theory. In order to highlight the natural structures related to generalized entropies, we use the central concepts of the modern mathematics.

The paper is organized as follows. In the Section 2, Subsection 2.1, we define

a special category related to measurable spaces (referred to as *MES*), enabling the introduction an *associated basic functional* Z_p (see the forthcoming Section for his exact definition). Both the Tsallis and Rényi entropies, as well as the distance in L^p spaces, may be expressed in terms of this functional. In the Subsection 2.2, we define the direct product of the objects in *MES* and we prove that the functional Z_p satisfies a *compatibility relation* with respect to this product i.e., it is *multiplicative*. This multiplicative property is equivalent to the additivity of the Rényi entropy. In the Subsection 2.3 we define the disjoint sum (or the coproduct) of the objects in *MES*, and we prove that the functional Z_p satisfies a *compatibility relation* with respect to coproduct i.e., it is *additive*. Note that this property is equivalent to one of the postulates characterizing the Rényi entropy. The prove that both product and coproduct possess a universal property and that the direct product and coproduct can also be defined for morphisms of the category *MES*, can be found in the Subsection 2.4. In the Subsection 2.5 we show that, by extending the category *MES* with the introduction of the *unit object* and the *null object*, the category *MES* became to a *monoidal category*.

Section 3 deals with the axiomatic characterization of the functional Z_p . We demonstrate that there exists a universal exponent p (the same for all the objects of the category) that characterizes completely the functional Z_p (hence, also the Tsallis or Rényi entropies) up to an arbitrary multiplicative factor.

Appendix 5.1 shows that the Rényi divergence can be expressed in terms of the Rényi entropy. The proof of the universality (with respect to all the objects of the category *MES*) of the exponent defining the Rényi or Tsallis entropies can be found in Appendix 5.2.

2 The category-theoretic properties related to Rényi and Tsallis entropies.

2.1 Definitions

Our definitions include as a particular case the original definition of the generalized entropies [5], [6] and [2]. Our basic construction that will play the role of the *object of the category MES* is derived from the well known concept of measurable space [16], [17]. Guided by statistical ideas, in order to take into account the *negligible sets* we specify also an sub-ideal of the σ -algebra of measurable sets. *The objects of the category MES* consist of triplets $M_X := (X, \mathcal{A}_X, \mathcal{N}_X)$ with X denoting the phase space (for instance, it is a symplectic manifold in the case of statistical physics or, in the case of elementary probability models, finite or denumerable set) and \mathcal{A}_X is the σ -algebra generated by a family of subsets of X , respectively. We also denote with $\mathcal{N}_X \subset \mathcal{A}_X$ an ideal of the σ -algebra \mathcal{A}_X having the meaning of *negligible sets*. Let us now postulate the *completeness property*. From $N \in \mathcal{N}_X$ and $N' \subset N$ results $N' \in \mathcal{N}_X$. *The morphisms of the category MES* with the source M_X and range M_Y are the measurable maps Φ from X to Y , which are *nonsingular* i.e., such that $\Phi^{-1}(\mathcal{N}_Y) \subset \mathcal{N}_X$. From the

completeness property results the *ideal property* i.e., if $N \in \mathcal{N}_X$ and $A \in \mathcal{A}_X$ then $A \cap N \in \mathcal{N}_X$. Note that it is possible that \mathcal{N}_X contains only the empty set (as, for example, in the case of atomic spaces).

Remark 1 *At first sight it would be more natural to consider the objects as measure space triplet $(X, \mathcal{A}_X, \mu_X)$ containing the measure μ_X , and the morphisms as the measure preserving transformations. However, in this case we cannot define direct product or coproduct having universal property.*

We denote with $C(M_X)$, or with $C(X, \mathcal{A}_X, \mathcal{N}_X)$, the cone with all σ -finite positive measures over $(X, \mathcal{A}_X, \mathcal{N}_X)$ that are compatible with \mathcal{N}_X (i.e., $\mu \in C(X, \mathcal{A}_X, \mathcal{N}_X)$ iff for all $N \in \mathcal{N}_X$ we have $\mu(N) = 0$). For a given $\mu_X \in C(X, \mathcal{A}_X, \mathcal{N}_X)$ and $p > 0$, we denote with $L^p(M_X, \mu_X)$ the Banach space ($p \geq 1$) or the Fréchet space ($0 < p < 1$) of functions $f_X : X \rightarrow \mathbb{R}$ that are measurable modulo \mathcal{N}_X and have finite norm (pseudo norm, respectively): more precisely, $\int_X |f_X(x)|^p d\mu_X(x) < \infty$. In the sequel, we shall denote

$$Z_p(M_X, \mu_X, \rho_X) := \int_X \rho_X(x)^p d\mu_X(x) \quad (1)$$

for some non-negative density $\rho_X \in L^p(M_X, \mu_X)$. The generalized entropies are defined for probability density functions (PDF) satisfying the conditions

$$\rho_X \in L^1(M_X, \mu_X) \cap L^p(M_X, \mu_X); \quad (2)$$

$$\int_X \rho_X(x) d\mu_X(x) = 1 \quad (3)$$

where $p > 0$ and $p \neq 1$. The probability $P(A)$ can be represented by PDF as follows

$$P(A) = \int_A \rho_X(x) d\mu_X(x); \quad (4)$$

$$A \subset X; A \in \mathcal{A}_X; \mu_X \in C(M_X) \quad (5)$$

In this framework, for a given measurable space $M_X := (X, \mathcal{A}_X, \mathcal{N}_X)$ and measure $\mu_X \in C(M_X)$, the classical Boltzmann-Gibbs-Shannon entropy functional is given by

$$S_{cl}[M_X, \mu_X, \rho_X] = - \int_X \rho_X(x) \log [\rho_X(x)] d\mu_X(x) \quad (6)$$

For a given measurable space M_X , the generalizations of the A. Rényi [2] and C. Tsallis [5], [6] entropies, involves the functional $Z_p(M_X, \mu_X, \rho_X)$ given by Eq.(1). The functional Z_p is related to the norm of the density ρ in the Banach space for $p \geq 1$ [17], and to the pseudo-norm $N_p[\rho]$ for $0 < p < 1$ [16] [18],

through the obvious relations

$$\|\rho_X\|_p = \left[\int_X [\rho_X(x)]^p d\mu_X(x) \right]^{\frac{1}{p}} ; p \geq 1 \quad (7)$$

$$N_p[\rho_X] = \int_{\Omega} [\rho_X(x)]^p d\mu_X(x); 0 < p \leq 1 \quad (8)$$

These relations give the geometrical interpretation of the generalized entropies (for further information Refs to [13]).

Remark 2 *The study of the generalized entropies helps us to better understand the classical entropy. For $p \geq 1$, the functional $\|\rho_X\|_p$ is the classical L^p norm, and for $0 < p < 1$ the functional $N_p[\rho_X]$ is the exotic L^p -norm [18]. For $p > 1$ the L^p spaces are reflexive, the Maxent problem is equivalent to the minimal L^p distance problem with restrictions [13], or to the minimal $Z_p(M_X, \mu_X, \rho_X)$. For $0 < p < 1$, the L^p spaces has, in general, trivial duals, the Maxent problem is equivalent to the maximal L^p distance or the maximal $Z_p(M_X, \mu_X, \rho_X)$ (see [13]). The case $p = 1$, which corresponds to the classical Shannon entropy, is just the border point between two radically different functional-analytic properties.*

The corresponding generalized entropy $S_{R,p}$, proposed by A. Rényi [2], and the entropy, $S_{T,p}$, proposed by C. Tsallis [5], [6] are given by

$$S_{R,p}[M_X, \mu_X, \rho_X] = \frac{1}{1-p} \log Z_p(M_X, \mu_X, \rho_X) \quad (9)$$

$$S_{T,p}[M_X, \mu_X, \rho_X] = \frac{1}{1-p} [1 - Z_p(M_X, \mu_X, \rho_X)] \quad (10)$$

Consider now a measure space $N = (\Omega, \mathcal{A}, n)$ with σ -finite measure n , and let us denote with $P(x)$, $Q(x)$ two probability densities:

$$\int_{\Omega} P(x) dn(\mathbf{x}) = \int_{\Omega} Q(x) dn(\mathbf{x}) = 1$$

Note that the Rényi divergence [2], [12]

$$D_p(P||Q) = \frac{1}{p-1} \log \int_{\Omega} P^p Q^{1-p} dn(\mathbf{x}) \quad (11)$$

is related to the Rényi entropies, Eq.(9), by Eq.(48) (see Appendix 5.1). Note that when X is a finite or denumerable set, if we denote with p_k the probabilities of element $x_k \in X$, the measure μ_X is the counting measure on the space X (equal to the number of elements in a subset), and the family of null sets

$\mathcal{N}_X = \{\emptyset\}$ then, from the previous Eqs.(9, 10, 1) we get the original definitions from Ref.[2], [5],[6]

$$S_{R,q}[M_X, \mu_X, \rho_X] = \frac{1}{1-q} \log \sum_k p_k^q \quad (12)$$

$$S_{T,q}[M_X, \mu_X, \rho_X] = \frac{1}{1-p} \left[1 - \sum_k p_k^q \right] \quad (13)$$

$$Z_q(M_X, \mu_X, \rho_X) = \sum_k p_k^q \quad (14)$$

Remark that, in this particular case, $S_{T,q}[M_X, \mu_X, \rho_X]$, as well as $Z_q(M_X, \mu_X, \rho_X)$, are Lesche stable [19]. Note that, from Eqs (6, 9 and 10), results

$$\lim_{p \rightarrow 1} S_{T,q}[M_X, \mu_X, \rho_X] = \lim_{p \rightarrow 1} S_{R,q}[M_X, \mu_X, \rho_X] = S_{cl}[M_X, \mu_X, \rho_X] \quad (15)$$

2.2 Direct product of measurable spaces and the multiplicative property of $Z_p[M_X, \mu_X, \rho_X]$

In the framework of the our formalism, the multiplicative property is the counterpart of the *Postulate 4* in the Rényi theory[2]. In the following we overload the tensor product notation " \otimes "; its meaning results from the nature of the operand. Denote the direct product of two measurable spaces $M_X = (X, \mathcal{A}_X, \mathcal{N}_X)$ and $M_Y = (Y, \mathcal{A}_Y, \mathcal{N}_Y)$ by $M_X \otimes M_Y$, defined as follows

$$M_X \otimes M_Y = (X \times Y, \mathcal{A}_X \otimes \mathcal{A}_Y, \mathcal{N}_{X \otimes Y}) \quad (16)$$

Here $X \times Y$ is the *Cartesian product* of the phase spaces X and Y , while the σ -algebra $\mathcal{A}_X \otimes \mathcal{A}_Y$ is the smallest σ -algebra such that it contains all of the elements of the Cartesian product $\mathcal{A}_X \times \mathcal{A}_Y$. The *null set ideal* $\mathcal{N}_{X \otimes Y} \subset \mathcal{A}_X \otimes \mathcal{A}_Y$ is generated by the family $(\mathcal{A}_X \otimes \mathcal{N}_Y) \cup (\mathcal{N}_X \otimes \mathcal{A}_Y)$. Note that if $\mu_X \in C[M_X]$ and $\mu_Y \in C[M_Y]$ then their direct product satisfies the condition $\mu_X \otimes \mu_Y \in C[M_X \otimes M_Y]$ (we denote it also by the same symbol). The measure $\mu_X \otimes \mu_Y$ acting on $(\mathcal{A}_X \otimes \mathcal{A}_Y)/\mathcal{N}_{X \otimes Y}$ are defined by extension by denumerable additivity, starting from the product subsets:

$$(\mu_X \otimes \mu_Y)(A_X \times A_Y) = \mu_X(A_X) \mu_Y(A_Y) \quad (17)$$

$$A_X \in \mathcal{A}_X; A_Y \in \mathcal{A}_Y \quad (18)$$

Consider now the measures $\mu_X \in C(M_X)$, $\mu_Y \in C(M_Y)$, and the densities $\rho_X \in L^p(M_X, d\mu_X) \cap L^1(M_X, d\mu_X)$ and $\rho_Y \in L^p(M_Y, d\mu_Y) \cap L^1(M_Y, d\mu_Y)$. The following function is also denoted with the same symbol

$$\rho_X \otimes \rho_Y \in L^p(M_X \times M_Y, \mu_X \otimes \mu_Y) \cap L^1(M_X \times M_Y, \mu_X \otimes \mu_Y) \quad (19)$$

with

$$(\rho_X \otimes \rho_Y)(x, y) = \rho_X(x) \rho_Y(y) \quad (20)$$

$$x \in X; y \in Y \quad (21)$$

We have the following basic proposition

Proposition 3 *Let ρ_X , ρ_Y are normalized PDF*

$$\int_X \rho_X(x) d\mu_X(x) = \int_Y \rho_Y(y) d\mu_Y(y) = \mathbf{1}; \quad \rho_X \geq 0; \quad \rho_Y \geq 0 \quad (22)$$

Then we have

$$Z_p[M_X \otimes M_Y, \mu_X \otimes \mu_Y, \rho_X \otimes \rho_Y] = Z_p[M_X, \mu_X, \rho_X] Z_p[M_Y, \mu_Y, \rho_Y] \quad (23)$$

$$S_{R,p}[M_X \otimes M_Y, \mu_X \otimes \mu_Y, \rho_X \otimes \rho_Y] = S_{R,p}[M_X, \mu_X, \rho_X] + S_{R,p}[M_Y, \mu_Y, \rho_Y] \quad (24)$$

The validity of this statement follows directly from the definitions of the direct product, the Rényi entropy and the functional Z_p .

2.3 Coproduct of measurable spaces and the additivity of the functional $Z_p[M_X, \mu_X, \rho_X]$

Let us study now the property encoded in the *Postulate 5'* related to the Rényi entropy theory (Ref. [2]), transcribed in the measure theoretic and category language and re-expressed in the term of the functional $Z_p[M_X, \mu_X, \rho_X]$. Also in this case, we overload the notation \sqcup , for the disjoint sum from the set theory. Its precise meaning will be clear from the nature of the operands. In the following we investigate the *functorial properties*, related to *Postulate 5'*, of the functional $Z_p[M_X, \mu_X, \rho]$, in analogy to Proposition 3. To this end we introduce the following

Definition 4 *The coproduct of measurable spaces $M_X = (X, \mathcal{A}_X, \mathcal{N}_X)$ and $M_Y = (Y, \mathcal{A}_Y, \mathcal{N}_Y)$ will be denoted by $M_X \sqcup M_Y$ and have the following structure*

$$M_X \sqcup M_Y = (X \sqcup Y, \mathcal{A}_X \sqcup \mathcal{A}_Y, \mathcal{N}_X \sqcup \mathcal{N}_Y) \quad (25)$$

Here, $X \sqcup Y$ is the disjoint sum of the sets X and Y , and $\mathcal{A}_X \sqcup \mathcal{A}_Y$ is the smallest σ -algebra that contains all of the sets of the form $A_1 \sqcup A_2$, with $A_1 \in \mathcal{A}_X$ and $A_2 \in \mathcal{A}_Y$, respectively. Moreover, the new null set ideal $\mathcal{N}_X \sqcup \mathcal{N}_Y$ is the smallest σ -algebra generated by the family $N_1 \sqcup N_2$ with $N_1 \in \mathcal{N}_X$ and $N_2 \in \mathcal{N}_Y$. Let the measures $\mu_X \in C(M_X)$, $\mu_Y \in C(M_Y)$ and the weights $w_1 \geq 0$, $w_2 \geq 0$ and $w_1 + w_2 = 1$. The measure $\mu := w_1 \mu_X \sqcup w_2 \mu_Y$ acts on the σ -algebra $\mathcal{A}_X \sqcup \mathcal{A}_Y$ and it is defined uniquely as the continuation by denumerable additivity from the property

$$\mu(A_1) = w_1 \mu_X(A_1); \quad A_1 \in \mathcal{A}_X \quad (26)$$

$$\mu(A_2) = w_2 \mu_Y(A_2); \quad A_2 \in \mathcal{A}_Y \quad (27)$$

Let $\rho_X \in L^p(M_X, d\mu_X) \cap L^1(M_X, d\mu_X)$ and $\rho_Y \in L^p(M_Y, d\mu_Y) \cap L^1(M_Y, d\mu_Y)$. We define the function $\rho := \rho_X \sqcup \rho_Y \in L^p(M_X \sqcup M_Y, w_1 \mu_X \sqcup w_2 \mu_Y) \cap L^1(M_X \sqcup M_Y, w_1 \mu_X \sqcup w_2 \mu_Y)$ as follows

$$\rho(x) = \rho_X(x); \quad \text{if } x \in X$$

$$\rho(x) = \rho_Y(x); \quad \text{if } x \in Y$$

We restrict our definition of coproduct to finite terms. An example of (denumerable infinite) coproduct is the grand canonical ensemble.

Remark 5 *If $\rho_X d\mu_X$ and $\rho_Y d\mu_Y$ are probability measures, then the measure $[\rho_1 \sqcup \rho_2](x)[w_1 d\mu_X \sqcup w_2 d\mu_Y]$ is a probability measure if $w_1 + w_2 = 1$.*

From the previous definition of the direct sum and the functional $Z_p[M_X, \mu_X, \rho_X]$ the following obvious proposition results

Proposition 6 *The reformulation of the Postulate 5' (Ref. [2]) reads: the functional $Z_p[M_X, \mu_X, \rho_X]$ is additive with respect to the direct sum of measurable spaces*

$$Z_p[M_X \sqcup M_Y, w_1 \mu_X \sqcup w_2 \mu_Y, \rho_X \sqcup \rho_Y] = w_1 Z_p[M_X, \mu_X, \rho_X] + w_2 Z_p[M_Y, \mu_Y, \rho_Y] \quad (28)$$

2.4 Universal properties of the direct product and direct sum in the category of measurable spaces

In the following we prove that the basic binary operations on measurable spaces, the direct product and the direct sum, defined in the previous section, have universality properties in the category of measurable spaces MES .

Consider the direct product $M = M_X \otimes M_Y$ of measurable spaces $M_X = (X, \mathcal{A}_X, \mathcal{N}_X)$ and $M_Y = (Y, \mathcal{A}_Y, \mathcal{N}_Y)$. Observe that the canonical projections $p_X : X \times Y \rightarrow X$, $p_Y : X \times Y \rightarrow Y$, are measurable and induce the morphisms $\pi_X : M_X \otimes M_Y \rightarrow M_X$ and $\pi_Y : M_X \otimes M_Y \rightarrow M_Y$ between the objects of MES . We have the following

Proposition 7 *In the category MES the applications $\pi_X : M_X \otimes M_Y \rightarrow M_X$, $\pi_Y : M_X \otimes M_Y \rightarrow M_Y$, which are naturally induced by canonical projections $p_X : X \times Y \rightarrow X$ and $p_Y : X \times Y \rightarrow Y$, are morphisms.*

Proof. *The measurability of π_X is direct consequence of the fact that the canonical projection maps are measurable, in fact the measurability of the canonical projections is an alternative definition of the product of σ algebras. The nonsingularity property $p_X^{(-1)}(\mathcal{N}_X) \subset \mathcal{N}_{X \times Y}$ results from $p_X^{(-1)}(\mathcal{N}_X) = \mathcal{N}_X \times \mathcal{A}_Y \subset \mathcal{N}_{X \times Y}$. ■*

From the previous Proposition 7 results immediately the following Theorem

Theorem 8 *In the category MES , the direct product has the universal property. Let $M_X = (X, \mathcal{A}_X, \mathcal{N}_X)$, $M_Y = (Y, \mathcal{A}_Y, \mathcal{N}_Y)$ and $M = (Z, \mathcal{A}_Z, \mathcal{N}_Z)$ measurable spaces that are objects of the category MES , such that there exists morphisms $\phi_X \in \text{Hom}(M, M_X)$ and $\phi_Y \in \text{Hom}(M, M_Y)$. Then there exists an unique morphism $\theta \in \text{Hom}(M, M_X \otimes M_Y)$ such that*

$$\phi_X = \pi_X \circ \theta \quad (29)$$

$$\phi_Y = \pi_Y \circ \theta \quad (30)$$

where π_X, π_Y are the morphism defined in Proposition 7.

Proof. The morphism θ is induced by the application $T : Z \rightarrow X \times Y$ defined as $Z \ni z \rightarrow T(z) := (\phi_X(z), \phi_Y(z)) \in X \times Y$. and it is unique. In order to prove that θ is a morphism we have to prove that T is measurable and it is nonsingular. To prove that $T : Z \rightarrow X \times Y$ is measurable, we recall that it is sufficient to prove that, for all $A \in \mathcal{A}_X, B \in \mathcal{A}_Y$, we have the property $T^{(-1)}(A \times B) \in \mathcal{A}_Z$, a property resulting from the measurability of ϕ_X and ϕ_Y . Note that to prove the inclusion $T^{-1}(\mathcal{N}_{X \times Y}) \subset \mathcal{N}_Z$, it is sufficient to demonstrate for the generating subsets $T^{-1}(\mathcal{N}_X \times \mathcal{A}_Y) \subset \mathcal{N}_Z$ (which follows from the nonsingularity of ϕ_X and $T^{-1}(\mathcal{A}_X \times \mathcal{N}_Y) \subset \mathcal{N}_Z$) that this is the consequence of the nonsingularity of ϕ_Y . ■

In conclusion the direct product operation has the natural functorial property, so the multiplicative property Eq. (23) of the functional $Z_p(M_X, \mu_X, \rho_X)$ appears as an algebraic compatibility property. By simple reversal of the arrows, we are lead to the corresponding universality property of the coproduct in the category MES . We have the following obvious proposition

Proposition 9 *In the category MES , consider the objects M_X, M_Y . The applications $\iota_X : M_X \rightarrow M_X \sqcup M_Y$ and $\iota_Y : M_Y \rightarrow M_X \sqcup M_Y$, induced naturally by the canonical injections $i_X : X \rightarrow X \sqcup Y, i_Y : Y \rightarrow X \sqcup Y$, are morphism in the category MES .*

Proof. The injections i_X, i_Y are measurable. Suppose that $N_1 \sqcup N_2 \in \mathcal{N}_X \sqcup \mathcal{N}_Y$, with $N_1 \in \mathcal{N}_X, N_2 \in \mathcal{N}_Y$ (see Definition 4). Then, $i_X^{(-1)}(N_1 \sqcup N_2) = N_1, i_Y^{(-1)}(N_1 \sqcup N_2) = N_2$, so i_X and i_Y are nonsingular, which completes the proof that ι_X, ι_Y are morphisms in the category MES . ■

By reversing the arrows, in analogy to the Theorem 8, we obtain the following result

Theorem 10 *In the category MES the direct sum of the objects has the following universality property. Let denote with $M_X = (X, \mathcal{A}_X, \mathcal{N}_X), M_Y = (Y, \mathcal{A}_Y, \mathcal{N}_Y)$ and $M = (Z, \mathcal{A}_Z, \mathcal{N}_Z)$ measurable spaces that are objects of the category MES , such that there exists morphisms $\phi_X \in \text{Hom}(M_X, M)$ and $\phi_Y \in \text{Hom}(M_Y, M)$. Then, there exists an unique morphism $\gamma \in \text{Hom}(M_X \sqcup M_Y, M)$ such that*

$$\begin{aligned}\gamma \circ \iota_X &= \phi_X \\ \gamma \circ \iota_Y &= \phi_Y\end{aligned}$$

where ι_X, ι_Y are the morphisms defined in Proposition 9.

Proof. The morphism γ is induced by the map $g : X \sqcup Y \rightarrow Z$ defined as follows. If $x \in X$ then $g(x) := \phi_X(x) \in Z$, and in the case $x \in Y$, then $g(x) := \phi_Y(x) \in Z$. The measurability of the map g results from the measurability of ϕ_X and ϕ_Y . The inclusion $g^{(-1)}(\mathcal{N}_Z) \subset \mathcal{N}_X \sqcup \mathcal{N}_Y$ results from the nonsingularity of ϕ_X and ϕ_Y . ■

In conclusion, the direct sum operation has natural category theoretic properties. Hence, the additivity property Eq.(28) of the functional $Z_p(M_X, \mu_X, \rho_X)$ is not an artificial construction.

2.5 The monoidal categories associated to product and coproduct

We recall the following

Proposition 11 [15] *Let \mathcal{C} be a category such that for all objects $A, B \in \text{Ob}(\mathcal{C})$ exists their direct product $A \otimes B$, having the universal property. Then, there exists a covariant functor F from the product category to \mathcal{C} , $\mathcal{C} \times \mathcal{C} \rightarrow \mathcal{C}$, defined as follows. For the object (A, B) of $\mathcal{C} \times \mathcal{C}$, where A, B are objects of \mathcal{C} , we have*

$$F((A, B)) := A \otimes B$$

For the pair of morphisms $(u, v) \in \text{Hom}((A, B), (A', B'))$ with $u \in \text{Hom}(A, A')$, $v \in \text{Hom}(B, B')$, from the category $\mathcal{C} \times \mathcal{C}$ there exists an unique morphism w in the category \mathcal{C} , $w \in \text{Hom}(A \otimes B, A' \otimes B')$ uniquely fixed by the conditions

$$\begin{aligned} w &= F((u, v)) \\ p_{A'} \circ w &= u \circ p_A \\ p_{B'} \circ w &= v \circ p_B \end{aligned}$$

We denoted with p_A, p_B the projections from $\text{Hom}(A \otimes B, A)$, $\text{Hom}(A \otimes B, B)$, and $p_{A'}$ are $p_{B'}$ the projections from $\text{Hom}(A' \otimes B', A')$, $\text{Hom}(A' \otimes B', B')$. The map $(u, v) \rightarrow F((u, v))$ has the functorial property.

Let $(u, v) \in \text{Hom}((A, B), (A', B'))$ and $(u', v') \in \text{Hom}((A', B'), (A'', B''))$. Then,

$$F((u' \circ u, v' \circ v)) = F((u', v')) \circ F((u, v)) \in \text{Hom}(A \otimes B, A'' \otimes B'')$$

If in the category \mathcal{C} we have an unit object, then \mathcal{C} is a monoidal category.

Similarly, by duality arguments, we have the following result for the direct sum (coproduct)

Proposition 12 [15] *Let \mathcal{C} be a category such that for all objects A, B from $\text{Ob}(\mathcal{C})$ exists their direct sum $A \sqcup B$, having the universal property. Then, there exists a covariant functor G from the product category $\mathcal{C} \times \mathcal{C} \rightarrow \mathcal{C}$ defined as follows. For the object (A, B) of $\mathcal{C} \times \mathcal{C}$, where A, B are objects of \mathcal{C} we have*

$$G((A, B)) := A \sqcup B$$

For the pair of morphisms $(u, v) \in \text{Hom}((A, B), (A', B'))$, with $u \in \text{Hom}(A, A')$ and $v \in \text{Hom}(B, B')$, from the category $\mathcal{C} \times \mathcal{C}$ there exists an unique morphism w in the category \mathcal{C} , $w \in \text{Hom}(A \sqcup B, A' \sqcup B')$ uniquely fixed by the conditions

$$\begin{aligned} w &= G((u, v)) \\ w \circ i_A &= i_{A'} \circ u \\ w \circ i_B &= i_{B'} \circ v \end{aligned}$$

We denote with i_A, i_B the canonical injections from $\text{Hom}(A, A \sqcup B)$, $\text{Hom}(B, A \sqcup B)$, and with $i_{A'}, i_{B'}$ the injections from $\text{Hom}(A', A' \sqcup B')$, $\text{Hom}(B', A' \sqcup B')$. The association $(u, v) \rightarrow G((u, v))$ has the functorial property. Let $(u, v) \in \text{Hom}((A, B), (A', B'))$ and $(u', v') \in \text{Hom}((A', B'), (A'', B''))$ then,

$$G((u' \circ u, v' \circ v)) = G((u', v')) \circ G((u, v)) \in \text{Hom}(A \sqcup B, A'' \sqcup B'')$$

If in the category \mathcal{C} we have a null object then, \mathcal{C} is a monoidal category with respect to direct sum.

We emphasize that, despite the fact that the construction of the direct sum is dual to the direct product, from the previous proposition (12) the functor G is a covariant functor. In the category MES we have an unit object as well as the null object. The unit object is denoted with $M_1 := (1, \mathcal{A}_1, \mathcal{N}_1)$, where 1 is the one point set [15], \mathcal{A}_1 is the trivial σ -algebra consisting in the one point set 1, \emptyset , and $\mathcal{N}_1 = \{\emptyset\}$, respectively. The (more or less formal) null object M_0 , with respect to the direct sum, is the object generated by the empty set $M_0 := (\emptyset, \mathcal{A}_\emptyset, \mathcal{N}_\emptyset)$. So we have the following

Conclusion 13 *The category MES is a monoidal category both with respect to the product \otimes and the coproduct \sqcup .*

3 Axioms

We expose another approach, based on category theory, to the problem of the naturalness of the choice of the family of functions g_α used in the definition of the entropy [2]. We prove that this problem may be treated if we take into account the additivity and the multiplicative properties of the functional Z_p . We mention that a possible candidate for the generalization of the symmetry *Postulate 1* [2] is the requirement of invariance of the generalized entropy under measure preserving transformations. Recall that the group generated by finite permutations is the maximal *measure preserving group* with respect to the counting measure. The problem is that there are plenty of measures such that the measure preserving group is trivial (for instance, the atomic measure for 2 element set with $\mu(1) \neq \mu(2)$). To avoid this problem, we observe that *Postulate 1* and *Postulate 5'* in the original Rényi theory [2] can be generalized as follows. For a given measurable function $f(x)$ on the measured space M_X and $\mu \in C(M_X)$, let us define

$$m_f(M_X, \mu, t) = \mu[(x|x \in M_X \text{ \& } f(x) \leq t)] \quad (31)$$

Note that $m_f(M_X, \mu, t)$ is invariant under measure preserving transformations. In addition

$$Z_p[M_X, \mu_X, \rho] = \int_0^\infty t^p dm_\rho(M_X, \mu, t) \quad (32)$$

Then, the *Postulate 1* (the symmetry property) and *Postulate 5'* (the additivity property expressed in Proposition 6) can be generalized as follows. *Postulate 1* & *Postulate 5'*

$$Z_p[M_X, \mu_X, \rho_X] = \int_0^\infty h_X(t) dm_{\rho_X}(M_X, \mu, t) = \int_X h_X[\rho_X(x)] d\mu_X(x) \quad (33)$$

$$h_X(x) > 0; \text{ if } x > 0 \quad (34)$$

for some Borel measurable function $h_X(t)$ with

$$h_X(0) = 0 \quad (35)$$

The last requirement result by considering the case when the support of ρ_X is concentrated on a proper subset of X and by using Eq.(28) . The generalization of the *Postulate 2* (the continuity property) is straightforward. Be $h(x)$ continuos and $\rho_X \in L^1(M_X, \mu_X)$, we get

$$h_X[\rho(x)] \in L^1(M_X, \mu_X) \quad (36)$$

In our settings, the analog of the *Postulate 4* (the additivity property) [2] is the multiplicative property given by Eq.(23) and Proposition 3. By using Eqs.(23, 33, 35 and 36), and by continuity of the functions h_{XY} , h_X , h_Y for all $x, y \geq 0$, we obtain the following functional equation (valid almost everywhere)

$$h_{XY}(x y) = h_X(x)h_Y(y); \quad x, y \in \mathbb{R} \quad (37)$$

By arguments similar to the proof of the *uniqueness*, from Theorem 2 [2]), we get Eq. (32) (for details see the Appendix 5.2): there exists an universal family of functions, independent of X , parametrized by the positive parameter p such that

$$h_X(x) = x^p C_X \quad (38)$$

$$h_Y(y) = y^p C_Y \quad (39)$$

$$h_{XY}(z) = z^p C_X C_Y \quad (40)$$

4 Summary and conclusions

We proved that the most natural setting for treating the axiomatic approach to the study of definitions of measures of information or uncertainty, is the formalism of the category theory, that was invented for the most difficult, apparently contradictory aspects of the foundation of mathematics. In this respect we introduced a category of measurable spaces *MES*. We proved that in the category *MES* exist the direct product and the direct sum, having universal properties. We proved that the functional $Z_p(M_X, \mu_X, \rho_X)$ defined in Eq.(1), which appears in the definition of both Rényi and Tsallis entropies, has algebraic compatibility

properties with respect to direct product and direct sum, as shown in Eqs (23) and (28).

The main conclusions may be summarized as follows.

- (1) The natural measure of the quantity of information is the family of functionals $Z_p(M_X, \mu_X, \rho_X)$ given by Eq. (1), (defined in the Fréchet space for $0 < p < 1$, and in the Banach space for $p > 1$), and the classical Shannon entropy by Eq. (6);
- (2) The category MES is the natural framework for treating the problems related to the measure of the information, in particular in reformulating the Rényi axioms;
- (3) The category MES is a monoidal category with respect to direct product and coproduct and the functional $Z_p(M_X, \mu_X, \rho_X)$ has natural *compatibility properties* with respect to the product (it is multiplicative) and the coproduct (it is additive);
- (4) Up to a multiplicative constant, it is possible to recover the exact form of the functional $Z_p(M_X, \mu_X, \rho_X)$ defining the generalized entropies from a system of axioms that generalize the ones adopted by Rényi [2].

Acknowledgement 14 *The authors are grateful to Prof. M. Van Schoor and Dr D. Van Eester from Royal Military School, Brussels. György Steinbrecher is indepted to Prof. C. P. Niculescu from Mathematics Department, University of Craiova, Romania, and S. Barasch for discussions on category theory. Giorgio Sonnino is also grateful to Prof. P. Nardone and Dr P. Peeters of the Université Libre de Bruxelles (ULB) for useful discussions and suggestions.*

5 Appendix

5.1 Rényi divergence and entropy

Suppose to have a measurable space (Ω, \mathcal{A}, m) with a finite or σ -finite measure μ and a normalized PDF $\rho(x)$, i.e. $\int_{\Omega} \rho(x) d\mu(x) = 1$. Only in this Subsection we adopt the following definitions

$$U(\rho, d\mu, \alpha) := \int_{\Omega} [\rho(x)]^{\alpha} d\mu(x) \quad (41)$$

$$S_{R,\alpha}(\rho, d\mu) = \frac{1}{1-\alpha} \log U(\rho, d\mu, \alpha) \quad (42)$$

Consider now a measurable space $N = (\Omega, \mathcal{A}, n)$ with σ -finite measure n . We also denote with $P(x)$, $Q(x)$ two probability densities, satisfying the condition

$$\int_{\Omega} P(x) dn(x) = \int_{\Omega} Q(x) dn(x) = 1 \quad (43)$$

The Rényi divergence reads

$$D_p(P||Q) = \frac{1}{p-1} \log \int_{\Omega} P^p Q^{1-p} dn(x) \quad (44)$$

According to the notation Eq.(41) and normalization Eq.(43), we get

$$\int_{\Omega} Q^{1-p} dn(x) = U(Q, dn, 1-p) \quad (45)$$

and $P_1(x) := P(x)/U(Q, dn, 1-p)$ is normalized with respect to the measure $Q^{1-p}dn(\mathbf{x})$. Consequently, from Eq.(41), we find

$$\int_{\Omega} P^p Q^{1-p} dn(x) = U\left[\frac{P}{Z}, Q^{1-p}dn, p\right] Z^p \quad (46)$$

$$Z := U(Q, dn, 1-p) \quad (47)$$

By using the notation Eqs(42, 44-47) we obtain the following relation between the Rényi entropies and the divergences

$$D_p(P||Q) = -S_{R,p}\left[\frac{P}{Z}, Q^{1-p}dn\right] + \frac{p^2}{p-1}S_{R,1-p}[Q, dn] \quad (48)$$

5.2 Solution of the functional equation Eq.(37)

Using Eq. (34) with $\rho \geq 0$, we note that we can use the double logarithmic scale by performing the following change of variables

$$f_X(u) = \log h_X(\exp(u)) \quad (49)$$

$$f_Y(v) = \log h_Y(\exp(v)) \quad (50)$$

$$f_{XY}(z) = \log h_{XY}(\exp(z)) \quad (51)$$

Hence, Eq. (37) reads

$$f_{XY}(u+v) = f_X(u) + f_Y(v) \quad (52)$$

From Eq. (52), we obtain

$$f_{XY}(u+v) - f_{XY}(v) = f_X(u) - f_X(0) \quad (53)$$

$$[f_{XY}(u+v) - f_{XY}(v)] - [f_{XY}(u+0) - f_{XY}(0)] = 0 \quad (54)$$

$$f_{XY}(u+v) - f_{XY}(v) - f_{XY}(u) = f_{XY}(0) \quad (55)$$

The Eq.(56) admits the particular constant solution

$$F_{part}(z) \equiv f_{XY}(0) \quad (56)$$

The solution of corresponding homogenous equation

$$f_{XY}(u+v) - f_{XY}(v) - f_{XY}(u) = 0 \quad (57)$$

may be found by using again the continuity of the function $h_{XY}(\rho)$ (See also [20] I.3.1, page 8), i.e.,

$$f_{XY}(z) = pz \quad (58)$$

The general solution reads

$$f_{XY}(z) = f_{XY}(0) + pz$$

By using Eq. (52), we get the universal linear slope p

$$\begin{aligned} f_X(u) &= f_X(0) + pu \\ f_Y(v) &= f_Y(0) + pv \\ f_{XY}(0) &= f_X(0) + f_Y(0) \end{aligned}$$

and, by Eqs (49-51), up to undetermined multiplicative constants $C_X = \exp(f_X(0))$, $C_Y = \exp(f_Y(0))$, we find Eqs (38-40).

References

- [1] C. E. Shannon, Bell System Technical Journal, **27**, 379-423 & 623- 656 (1948).
- [2] A. Rényi (1961), *On measures of information and entropy, Proceedings of the fourth Berkeley Symposium on Mathematics, Statistics and Probability 1960*. pp. 547–561.
- [3] D. K. Fadeev, *Zum Begriff der Entropie einer endlichen Wahrscheinlichkeitsschernas*, (Arbeiten zur Informationstheorie I, Berlin, Deutscher Verlag der Wissenschaften, 1957).
- [4] A. Feinstein, *Foundations of Information Theory*, (McGraw-Hill, New York, 1958).
- [5] C. Tsallis, *Journal of Statistical Physics*, **52**, 479–487 (1988).
- [6] C. Tsallis, R. S. Mendes and A. R. Plastino, *Physica A* **261**, 534-554 (1998).
- [7] C. Tsallis, *Introduction to nonextensive statistical mechanics_ approaching a complex world*, (Springer, 2009).
- [8] M. Gell-Mann, C. Tsallis, *Nonextensive entropy- Interdisciplinary applications*, (Oxford University Press, USA, 2004).
- [9] G . Steinbrecher, B. Weyssow, *Phys. Rev. Lett.* **92** , 12503 (2004); T. L. Rhodes et al., *Phys. Lett. A* **253**, 181 (1999).
- [10] G. Steinbrecher, X. Garbet and B. Weyssow , "Large time behavior in random multiplicative processes", e- print arXiv:1007.0952v1 (2010).
- [11] T. Maszczyk and X. Duch, *Comparison of Shannon, Rényi and Tsallis Entropy used in Decision Trees*, *Lecture Notes in Computer Science*, **5097**, 643 (2008).

- [12] T. van Erven, P. Harremoës, "Rényi Divergence and Kullback-Leibler Divergence", e-print arXiv:1206.2459v2 (2014) .
- [13] G. Sonnino, G. Steinbrecher, Phys. Rev. **E 89**, 062106 (2014).
- [14] I. A. Bucur, *Introduction to the theory of categories and functors*, (Wiley, London, 1968).
- [15] Saunders Mac Lane, *Categories for the Working Mathematician*, (Springer, New York, 1997), second edition.
- [16] W. Rudin, *Real and Complex Analysis*, (McGraw Hill Inc. 1987) third edition.
- [17] M. Reed and B. Simon , *Methods of Modern Mathematical Physics*, (Academic Press , New York 1980), Vol. I. .
- [18] H. Luschgy and G. Pagès, Electronic Communication in Probability, **13**, 422-434 (2008).
- [19] B. Lesche, *Instabilities of the Rényi entropy*, Journal of Statistical physics, **27** , 419-422, (1982)
- [20] N. Dunford, J. T. Schwartz, *Linear Operators Part I: General Theory* , (J. Wiley, New York 1957).